Available online at www.sciencedirect.com

SCIENCE DIRECT®

Virology 323 (2004) 233–242

VIROLOGY

www.elsevier.com/locate/yviro

Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus*: a novel virus family, the *Globuloviridae*

Monika Häring,[a] Xu Peng,[b] Kim Brügger,[b] Reinhard Rachel,[a] Karl O. Stetter,[a] Roger A. Garrett,[b] and David Prangishvili[a,*]

[a] *Department of Microbiology, Archaea Centre, University of Regensburg, D-93053 Regensburg, Germany*

[b] *Danish Archaea Centre, Institute of Molecular Biology, Copenhagen University, DK-1307 Copenhagen K, Denmark*

Received 24 November 2003; returned to author for revision 8 January 2004; accepted 2 March 2004

Available online 27 April 2004

Abstract

A novel virus, termed *Pyrobaculum* spherical virus (PSV), is described that infects anaerobic hyperthermophilic archaea of the genera *Pyrobaculum* and *Thermoproteus*. Spherical enveloped virions, about 100 nm in diameter, contain a major multimeric 33-kDa protein and host-derived lipids. A viral envelope encases a superhelical nucleoprotein core containing linear double-stranded DNA. The PSV infection cycle does not cause lysis of host cells. The viral genome was sequenced and contains 28 337 bp. The genome is unique for known archaeal viruses in that none of the genes, including that encoding the major structural protein, show any significant sequence matches to genes in public sequence databases. Exceptionally for an archaeal double-stranded DNA virus, almost all the recognizable genes are located on one DNA strand. The ends of the genome consist of 190-bp inverted repeats that contain multiple copies of short direct repeats. The two DNA strands are probably covalently linked at their termini. On the basis of the unusual morphological and genomic properties of this DNA virus, we propose to assign PSV to a new viral family, the *Globuloviridae*.

Keywords: Virus; Archaea; Hyperthermophile; Linear genome; *Pyrobaculum*; *Thermoproteus*

Introduction

Viruses of the third domain of life, Archaea, have been characterized rudimentarily in contrast to their bacterial counterparts. Most viruses known to infect either haloarchaea or methanogens, of the kingdom Euryarchaeota, are typical head-and-tail phages with icosahedral heads and helical tails (reviewed by Dyall-Smith et al., 2003; Prangishvili and Zillig, 2001). In contrast, all known viruses infecting hyperthermophilic archaea of the kingdom Crenarchaeota exhibit unusual and diverse morphotypes. The latter have been classified into new viral families from these morphotypes: spindle-shaped *Fuselloviridae*, filamentous *Lipothrixviridae*, rod-shaped *Rudiviridae*, and droplet-shaped *Guttaviridae* (reviewed by Prangishvili, 2003; Prangishvili and Zillig, 2001; Prangishvili et al., 2001). The known hosts of these viruses are all members of the archaeal genera *Sulfolobus*, *Acidianus*, and *Thermoproteus*.

In addition to these cultivated hyperthermophilic viruses, diverse virus-like particles (VLPs) have been observed in sample enrichments from hot springs in Yellowstone National Park at temperatures above 80 °C (Rachel et al., 2002; Rice et al., 2001). Most were from acidic hot springs (pH 1.5–3), but one was from a neutral hot spring, Obsidian Pool (pH 6). Cells from acidic enrichments were invariably members of the genera *Sulfolobus* and *Acidianus*. Recently, we succeeded in isolating an *Acidianus* sp. from these enrichments that yielded a novel lipothrixvirus, AFV1 (Bettstetter et al., 2003).

VLPs with at least six different morphotypes were observed in an enrichment of a sample from Obsidian Pool. These included particles resembling head-and-tail phages of the *Siphoviridae* and *Podoviridae* families, and hyperthermophilic viruses of the *Rudiviridae* and *Lipothrixviridae*

* Corresponding author. Department of Microbiology, Archaea Centre, University of Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany. Fax: +49-941-943-2403.

E-mail address: david.prangishvili@biologie.uni-r.de (D. Prangishvili).

0042-6822/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2004.03.002

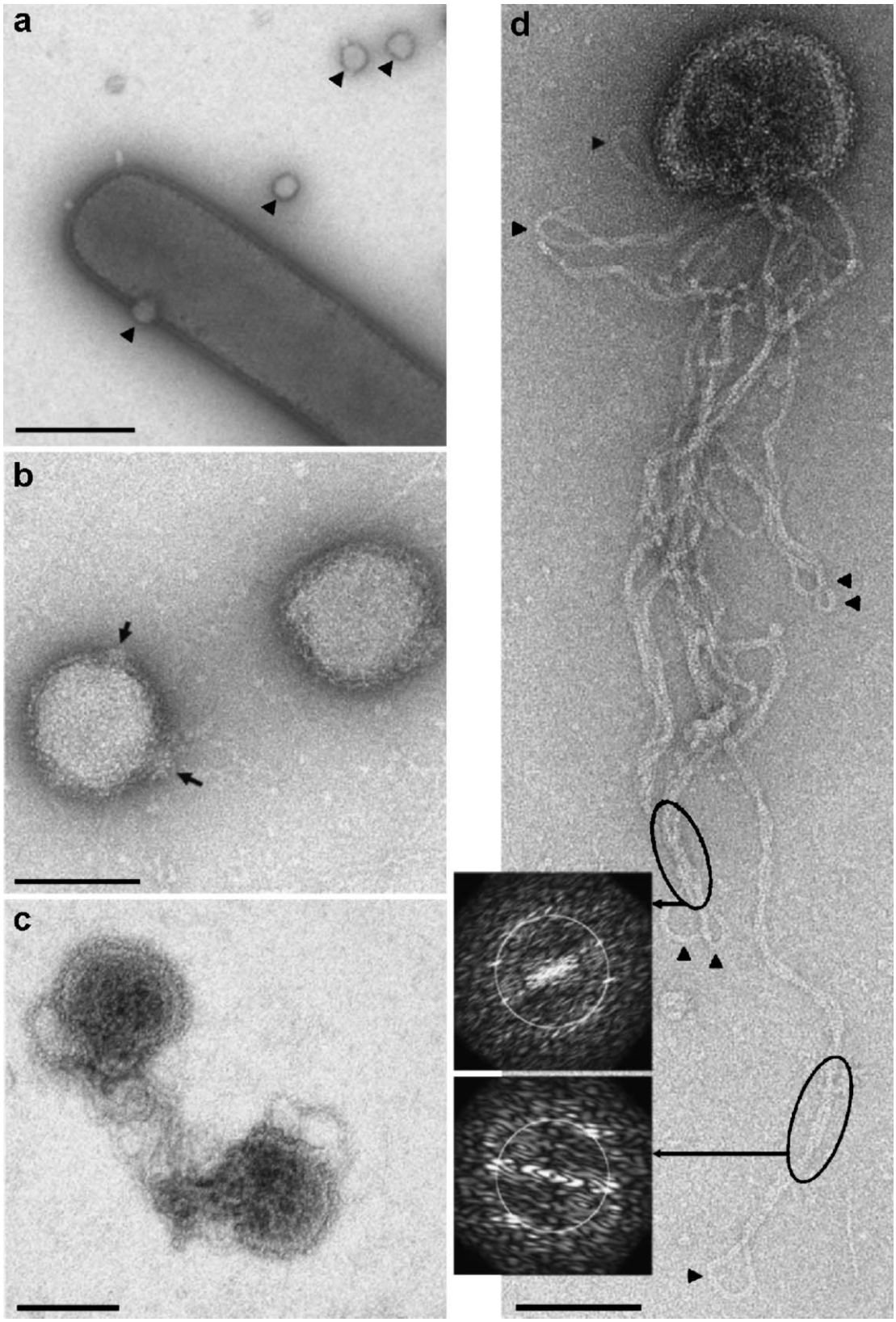

Fig. 1. Electron microscopy of *Pyrobaculum* sp. D11 and PSV virions negatively stained with 3% uranyl acetate. (a) Portion of a *Pyrobaculum* sp. D11 cell with four PSV virions marked by arrowheads; scale bar: 0.5 μm. (b) Two intact PSV virions, spherical protrusions are marked by arrows; scale bar: 0.1 μm. (c) Two disrupted PSV virions extruding disordered filamentous material; scale bar: 0.1 μm. (d) Disrupted PSV virion with extended filaments extruding from the particle. Several loops are marked by arrowheads, and two stretches enclosed by ellipsoids were analyzed by Fourier analysis as shown in the insets. The circle indicates a frequency of $(2.8\ \text{nm})^{-1}$; scale bar: 0.1 μm.

families, as well as spherical and unusual arrow-shaped pleomorphic VLPs (Rachel et al., 2002). Cells in the enrichment, which were potential hosts, were very diverse in their morphology and 16S rDNA sequences. The sequences obtained from the enriched samples indicated the presence of members of the bacterial genera *Thermus*, *Geothermobacterium*, and *Thermodesulfobacterium* and the archaeal genera *Thermophilum*, *Thermoproteus*, *Pyrobaculum*, *Thermospheaera*, and *Archaeoglobus*, as well as members of the archaeal Desulfurococcales, Korarchaeota, and Nanoarchaeota (Rachel et al., 2002). Here, we report the isolation of a novel spherical virus and a host strain *Pyrobaculum* sp. from this enrichment. The *Pyrobaculum* spherical virus (PSV) is the first virus to be described for the genus *Pyrobaculum*.

Results

Purification of VLPs

A sample from Obsidian Pool was enriched at 85 °C and pH 6 in a chemostat with a continuous flow of fresh medium and flushing with N_2 and CO_2 under strictly anaerobic conditions. Cells from a chemostat flow-through were removed by ultrafiltration and virus-like particles were concentrated and purified as described in Materials and methods. Isopycnic gradient centrifugation in CsCl yielded three sharp bluish-white opalescent bands with buoyant densities in the range of 1.305–1.338 g ml^{-1}. The bands contained the six types of VLP that were observed earlier (Rachel et al., 2002) in different proportions and were combined for further study. The morphology of the VLPs was unaltered after ultracentrifugation (data not shown).

Purification of virus–host systems

Twelve single isolates from the enrichment were obtained either using optical tweezers or by colony purification at 85 °C on Gelrite plates under anaerobic conditions. According to their 16S rDNA sequences, the isolates belong to the bacterial genera *Thermodesulfobacterium*, *Geothermobacterium*, and *Fervidobacterium* and to the archaeal genera *Pyrobaculum*, *Thermosphaera*, and *Thermofilum*.

Electron microscopy was employed to screen for virus-producing isolates. We tested for the presence of VLP growing cultures of all the single isolates, initially without success. Therefore, the sensitivity of the approach was increased by concentrating putative VLPs by ultrafiltration, as described in Materials and methods. However, again no VLPs were detected in the concentrated samples.

Single isolates were also screened for their ability to replicate any of the observed VLPs. A mixture of purified VLPs was added to the growing cultures of single isolates and the increase in particle numbers was determined by electron microscopy at different time intervals after "infection". In one isolate, cultured for 10 days (cell generation time about 24 h), we observed a high abundance of spherical particles, which was indicative of viral replication (Fig. 1a).

The host cells were assigned to a new species of the hyperthermophilic genus *Pyrobaculum*, termed *Pyrobaculum* sp. D11, of the crenarchaeal order Thermoproteales, from the 16S rDNA sequence (Fig. 2). The 6-fold symmetry of the S-layer, with a center-to-center distance between neighboring

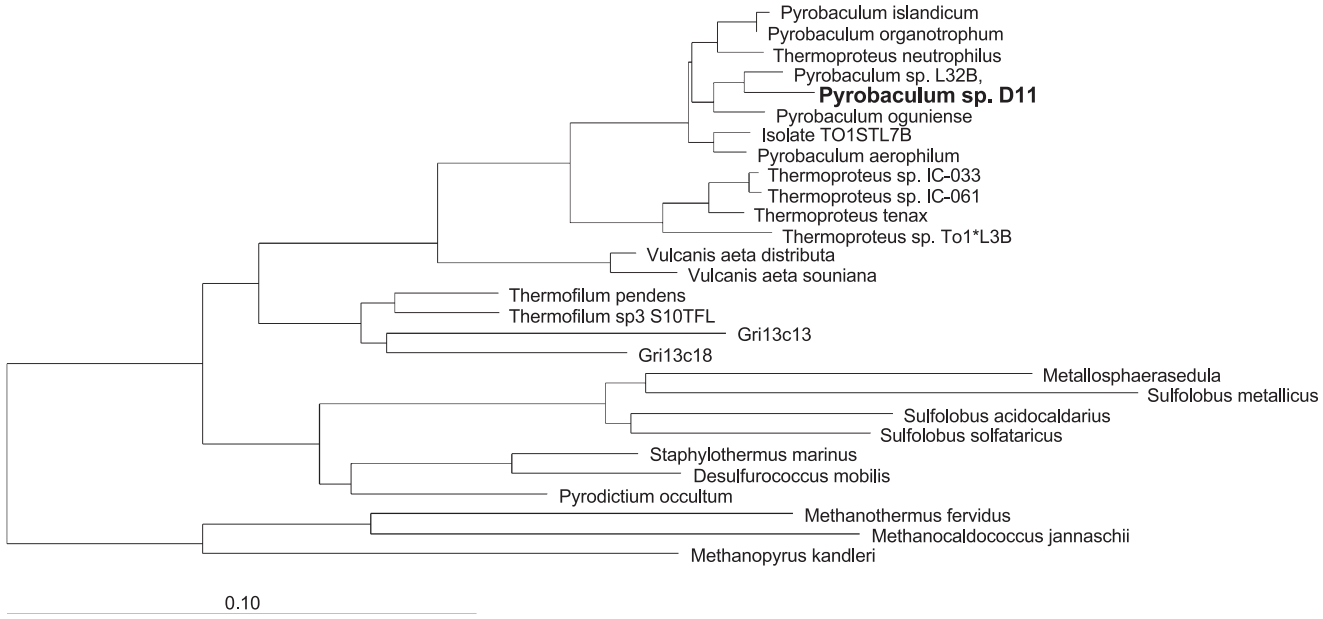

Fig. 2. Phylogenetic tree for Crenarchaeota determined by a neighbor-joining analysis of 16S rDNA sequences showing the position of *Pyrobaculum* sp. D11, the host of PSV, in bold type. The scale bar represents 0.10 fixed mutations per nucleotide position.

subunit complexes of about 30 nm (Fig. 1a), is typical for members of this order (Huber and Stetter, 2001). Although the host required strictly anaerobic growth conditions, virions of PSV were stable and remained infectious after storage for several months in the presence of oxygen. The new virus was termed *Pyrobaculum* spherical virus (PSV).

Virion morphology

The PSV virions were spherical with a diameter of about 100 nm (Fig. 1b). On most virions, we observed a variable number of spherical protrusions, about 15 nm in diameter (Fig. 1b). Although the structural integrity of most virions was unaffected by the purification procedures (including ultrafiltration and isopycnic ultracentrifugation), a small proportion of the purified virus particles was disrupted. From these disrupted particles, we observed the release of long filaments, 6 nm wide, probably containing nucleoprotein, either in a condensed form, with little detectable order (Fig. 1c), or in an extended form (Fig. 1d). A closer analysis revealed that the material consists of twin filaments with closed loops at the ends (Fig. 1d, arrowheads). Fourier analyses of small stretches of these filaments demonstrated a periodicity of about $(2.8\ \text{nm})^{-1}$, indicating a helical superstructure. These results suggest that PSV virions are enveloped and contain tightly packed nucleoprotein in a superhelical conformation.

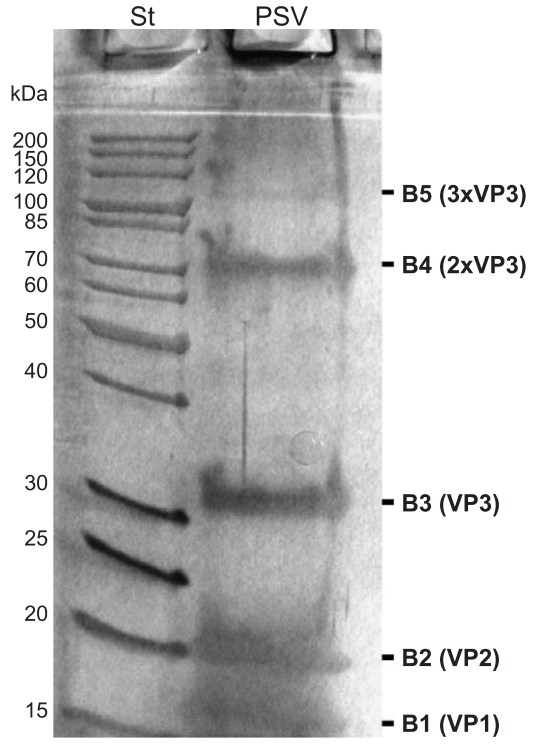

Fig. 3. SDS-PAGE of proteins extracted from the PSV virus stained with Coomassie Brilliant blue R-250. B1 to B5 indicate visible protein bands containing three viral proteins, VP1, VP2, and VP3. Bands B4 and B5 are considered further in the text. The molecular mass markers are indicated.

Fig. 4. Thin-layer chromatography of lipids extracted from PSV particles and uninfected cells of *Pyrobaculum* sp. D11. Ratio to front (RF) values of all spots are indicated. The solvent front, F, is indicated by a black line.

Virus–host relationships

The host range of PSV was examined by adding the purified virus to growing cultures of possible hosts, which included different strains of the closely related genera *Pyrobaculum* and *Thermoproteus*: *Pyrobaculum aerophilum*, *Pyrobaculum islandicum*, *Pyrobaculum organotrophum*, *Thermoproteus tenax*, and *Thermoproteus neutrophilus*. Virus propagation was tested by electron microscopy. The virus only replicated in one strain, sp. D11, of *Pyrobaculum* and in *T. tenax*.

No plaque assay was employed because the host strains could not be grown on plates as a lawn. An approximate

```
  1 CCTTGACCCA CCGCCAAGCC ACCCTCTACC AACGCCAAAC CACCTCCCCA
 51 CCAGACGCAA AAACACCCCC ACAAAAACGC AAAAACGCCC TTTACAAACC
101 GCCAAAAGAC GCCCAACCAA TGCCAACGCG AAACGCATAT ATCACAAAGA
151 CACTAAAGAC GCCAATGCCG CACGTCAACA TAATAGACGC acgccccgac
201 gctcccataa ctcaattcgc cctggagaat
```

Fig. 5. The 190-bp ITR sequence (in bold) at left terminus of the genome. The 5-bp repeats, A/GACGC (red) and CCACC (blue), and their close variants are colored. A 16-bp direct repeat is underlined.

estimation of the virus titer was obtained by counting the number of virions in meshes on copper grids in the electron microscope. As standards, we used preparations of the plaque-forming rudivirus SIRV2 of the hyperthermophilic *Sulfolobus islandicus* (Prangishvili et al., 1999) which had known titers.

Infection with PSV at an estimated moi of about 10 did not result in any significant increase in the 24-h generation time of *Pyrobaculum* sp. D11. Virus production was neither accompanied by a decrease in cell density, nor did we observe any significant amounts of cell debris in the culture. This indicated that the virus infection cycle did not, apparently, cause lysis of host cells. Viability of virus-infected cells was studied by a two-color fluorescence assay that allowed detection of cells with damaged membranes, as described in Materials and methods. Green fluorescent staining of all cells in the infected culture (not shown), at different stages of the infection cycle, was indicative of intact cell membranes and thus of a noncytocidal infectious cycle. After several successive transfers into fresh medium (dilution 1:1000) and continuous growth for several months, host cells still produced the virus.

Protein and lipid components

Analysis of the protein component of the virions by SDS-polyacrylamide gel electrophoresis yielded five protein bands, B1, B2, B3, B4, and B5, with molecular masses of about 16, 20, 33, 70, and 110 kDa, respectively (Fig. 3). The bands B2 and B3 were broad and extended, indicating that they might contain more than one component. However, N-terminal sequencing did not reveal any sequence heterogeneity of proteins eluted from these regions. The N-terminal sequence of the protein VP3 in band B3, ALVATTAVV-VYLQLSVVSVP, was identical to the N-terminal sequences of proteins in bands B4 and B5. The N-terminal sequence of the protein VP2 in band B2 was MLLPEGAGLLGLLA. No N-terminal sequence was obtained for protein VP1 in band B1 probably owing to N-terminal modification.

Lipids were extracted from viral particles by the chloroform-methanol method and analyzed by thin-layer chromatography, as described in Materials and methods. Four lipids were extracted from the virus particles (Fig. 4). It is unlikely that they derived from cell debris because the virus was purified by CsCl buoyant density centrifugation. Two of viral lipids had the same mobility as the two fastest migrating host lipids. However, the other two viral lipids differed in mobility from any host lipid (Fig. 4).

The genome and its organization

The genome was sequenced with a 6-fold coverage from a shot-gun library in a pUC18 vector with an average insert size of 2 kb. Sequence ambiguities were checked by primer walking on clones or directly on the viral DNA (see Materials and methods). No clones covering the extremities of the linear genome were obtained, as occurred for other archaeal linear viruses (Arnold et al., 2000; Bettstetter et al., 2003; Peng et al., 2001). Primers were designed to extend the sequencing at the ends by walking on amplified viral DNA. This procedure yielded an additional 693 bp of sequence at the left end and 807 bp of sequence at the right end. Analysis of these sequences revealed a 190-bp inverted terminal repeat (ITR) that contained multiple copies of the 5-bp repeats, A/GACGC and CCACC, and variants thereof, AATGC, AACAC, or CCACA (Fig. 5). A 16-bp direct repeat is also present in the ITRs (Fig. 5).

The termini of the genome were investigated by sequencing from two different primers annealed near the outer extremities of the ITRs. This yielded additional 20–30 nucleotides of sequence that stopped at the palindromic sequence CAAGGCCTTG. We conclude that this is the end of the viral genome. Moreover, we infer that the ends of the DNA strands are probably covalently closed such that a few nucleotides of sequence are read on the opposite strand (CCTTG), as occurred for the *Sulfolobus* rudiviruses SIRV1 and SIRV2 (Blum et al., 2001; Peng et al., unpublished data). The termination of readable sequence signals after the

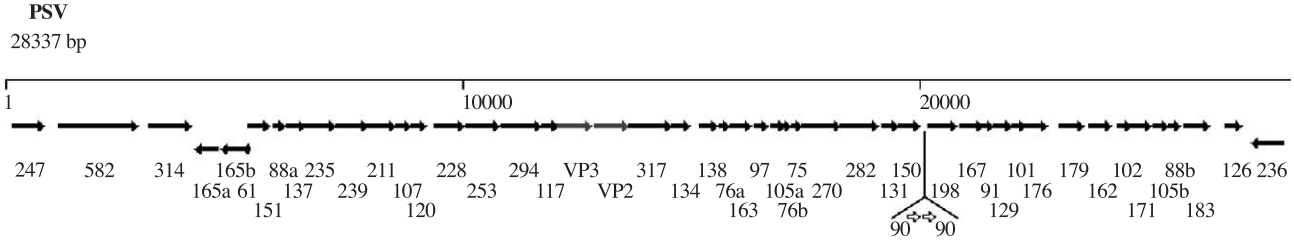

Fig. 6. Genome map of PSV showing positions and sizes of the ORFs listed in Table 1 and the direction of the gene transcripts. Only four putative ORFs, 61, 165a, 165, and 236, are encoded on the reverse strand. *VP2* and *VP3* encode virus proteins VP2 and VP3, respectively. The two inset ORF90s indicate the position of the triply repeated 241-bp sequence that occurs in about half of the clones (10 in total) sequenced from the PSV genome library.

CCTTG may be caused by a single-strand cut in DNA, as is found at the ends of rudiviral genomes (Blum et al., 2001). In the latter case, this cut was suggested to provide a free 3′-hydroxyl end for self-primed replication (Blum et al., 2001; Peng et al., 2001).

A few genomic heterogeneities were found during sequencing. These include a 241-bp sequence (positions 20 053–20 293) that was repeated perfectly three times in almost half of the 10 clones covering this region. The sequence encodes the C-terminal region of ORF 150 (Fig. 6). The presence of two additional copies of the sequence did not alter ORF 150, but they generated two identical copies of ORF 90 (Fig. 6), which is identical to the C-terminal region of ORF 150. The start codons of the two

Table 1
PSV ORFs and their putative regulatory sequences for transcription and translation

ORF	Promoter motif	S-D sequence	Terminator motif	Operons
Forward strand				
247	TTTACAAA (−21)		(+16) CTTCCTC–TTTTTTT	
582	TTAATAAA (−21)	GGCGT (−7)	(+1) GATTTTTCT	
314	TTATAA (−22)		(+3) CTTTTC–TATTTTTTCCT	
151	TTATTAA (−20)			
88a	AAATATAT (−19)			operon
137	TAAAATTTAAAAAT (−20)			
235				
239		GGCTG (−8)		
211				
107				
120			(+9) CTTTTTTCT	
228	ATTTATAT (−21)			operon
253	AATTATAAATA (−13)			
294	TTATTTAA (−20)			
117				
254 (VP3)	ATTATA (−26)		(+3) CTCCCTTTTTTCTCT	
246 (VP2)	AAATAATTTTTTAATAA (−19)			operon
317		GGTGG (−8)		
134	ATTTTTATATT (−20)		(+10) CTTTTTTCTTCTTT	
138			(+7) TTTTTTAA	
76a	TTTTTTAA (−20)			operon
163			(+3) TTCTT–TTTCTTTTTGA	
97	AAAAATTTAAATA (−21)		(+14) TTTTTTCCT	
105a	ATTTATATA (−20)			operon
76b	TTTAAATA (−19)			
75				
270	TTTAAAT (−21)			
282	TTATTTT (−28)		(+3) TTTTTTCTA	
131	TAAAATTTTTTAA (−21)			operon
150		GAGGG (−6)		
198	TTAAAAA (−20)		(+6) CTTTTTTG	
167	TAAAATTTAAAAAT (−20)			operon
91		CGGGG (−6)		
129		GGTGG (−9)		
101		GGGGG (−8)		
176		CGTGG (−7)	(+6) CTTTTC–TTTTTTCTTCTT	
179	TTTAAAAAT (−20)		(+9) TTTGCTTTT–CTTTTTTGTTTC	
162	TAAAATTTAAAAAT (−20)		(+13) TCTTTTT–CTTTTTTCCTTT	
102	TAAAATTTAAAAAT (−21)			operon
171			(+22) CTATTTCTTTTT	
105b	TTTTTA (−16)			operon
88b			(+3) TTGCTTTT	
183	TTTAAAA (−22)		(+17) TCTTTTTT–TCCTCTTTTTC	
126	TTTAAAAT (−20)		(+17) CTTTTTTCC–GCTTTTTT	
Reverse strand				
165a	AAAATTTATAAA (−21)		(+6) CCTTTTTT	
165b	ATTAAAA (−21)			operon
61		CGTGG (−6)		
236	ATAATTTAAT (−22)		(+22) TTTTCCC	

Numbers in parenthesis show positions upstream of the start codon for putative promoters and Shine–Dalgarno sequences, and downstream of the stop codon for putative terminators. Organization of genes in putative operons is indicated.

identical ORF90 are within the 3′ end of the first and the second copy of the repeat, respectively, and stop codons are within second and third copies, respectively. The two ORF90 are overlapping and not in the same frame. ORF90 did not yield any matches in a BLASTP search of public databases. No smaller direct or inverted repeats were found associated with the repeat sequence and the mechanism of DNA recombination is unclear. A total of 34 point mutations were detected in overlapping clones, 19 of which are randomly distributed in the ORFs with a strong preference for codon positions 1 and 2 (17 in total). Twelve of the remainder were concentrated in a 41-bp region between ORF176 and ORF179 (positions 23250–23290) (Fig. 6).

An ORF map was constructed by initially selecting for the largest possible gene using all three start codons (AUG, GUG, and TTG), and then examining each putative ORF for evidence of promoter and terminator motifs, and Shine–Dalgarno sequences using the experience gained from analyzing chromosomes and other genetic elements of crenarchaea (Reiter et al., 1988; Tolstrup et al., 2000). When their presence supported the existence of a shorter coding region, then this was selected. A total of 48 putative ORFs were identified and are presented on the genome map in Fig. 6. The numbers used to identify ORFs refer to the number of amino acid codons in the predicted proteins. Most ORFs carry recognizable regulatory sequence motifs for transcription and translation that are listed in Table 1.

Thirty ORFs are preceded by promoter box A motifs in the region −18 to −30 from the start codon (Table 1), suggesting that they are transcribed as mRNAs that are either leaderless or carry very short leaders, which lack Shine–Dalgarno motifs. This phenomenon is common for genes encoded in the chromosome of *P. aerophilum* (Slupska et al., 2001) and some other crenarchaea (Tolstrup et al., 2000) where, generally, single genes, or the first gene of an operon, are leaderless whereas genes within operons are preceded by Shine–Dalgarno sequences. In the PSV genome, most single genes and the first genes of operons follow this pattern. However, although eight of the nine putative Shine–Dalgarno sequences fall internally within operons, several genes within putative operons follow closely on from other genes but lack Shine–Dalgarno motifs (Table 1). The downstream T-rich sequences, which are likely to be transcriptional terminators (Reiter et al., 1988), generally follow either a single gene or the final gene of an operon (Table 1).

Exceptional for an archaeal double-stranded DNA virus is the concentration of the genes on one DNA strand (Fig. 6). There is only a cluster of three genes (positions 3700 and 5000), and ORF236 near the right hand terminus, that is present on the complementary strand (Fig. 6; Table 1).

Two genes were identified that encode the N-terminal sequences, MLLPEGAGLLGLLA and ALVATTAVVVYLQLSVVSVP, respectively (see above), of viral proteins VP2 and VP3. However, no genes were found that encoded proteins with molecular masses of 70 and 110 kDa (bands B4 and B5 in Fig. 3) with N-terminal sequences identical to that of VP3.

None of the ORFs yielded any significant matches with either nucleotide or protein sequences in public databases. The best matches obtained, for any of the ORFs, were to three smaller sections of larger hypothetical ORFs in the chromosome of *P. aerophilum* (e^{-5} to e^{-9}) that belongs to the same genus as the host *Pyrobaculum* sp. D11.

Discussion

Electron microscopy studies have demonstrated that spherical PSV virions consist of an envelope surrounding a nucleocapsid with a helical symmetry and width of about 6 nm. This structure is very unusual: in the viral world, no other enveloped DNA virus is known with a helical nucleocapsid (van Regenmortel et al., 2000). Enveloped, nonfilamentous DNA viruses, from archaea or bacteria, generally contain tightly packed coiled DNA. Superhelical arrays of nucleoprotein are more characteristic of enveloped eukaryal viruses, including paramyxoviruses, which have RNA genomes (van Regenmortel et al., 2000).

PSV is the first virus to be described for the hyperthermophilic archaeal genus *Pyrobaculum*. The virus replicates in both *Pyrobaculum sp*. D11, a newly isolated species of the genus, and in *T. tenax*, a strain of the very closely related genus *Thermoproteus*. The virus infection cycle apparently does not cause lysis of host cells. Very likely, similar to other hyperthermophilic archaeal viruses, PSV exists in its host in a carrier state. Such a strategy of virus–host relationship may enable hyperthermophilic viruses to minimize direct exposure to the extreme and hostile environmental conditions.

SDS-polyacrylamide gel electrophoresis of virus particles revealed five protein bands with molecular masses of about 16, 20, 33, 70, and 110 kDa. The three largest proteins yielded identical N-terminal sequences. Only one gene in the viral genome encodes such a sequence. The putative protein has 254 amino acid residues and an inferred molecular mass of about 28 kDa, similar to the estimated size of protein VP3 (Fig. 3), the smallest of the three proteins with identical N-termini. The results suggest that the two largest proteins are a dimer and trimer (or tetramer) of VP3. The latter protein contains no cysteine residues; however, it contains more than two-thirds hydrophobic amino-acid residues, about 15% of which are aromatic. This may contribute to the proteins' extreme thermostability and render it resistant to denaturation during SDS gel electrophoresis.

From virions of PSV could be extracted lipids by chloroform–methanol treatment. Thin-layer chromatography revealed a pattern of bands less complex than that of host lipids. The two fastest migrating lipids in the two patterns were identical in mobility. However, other viral lipids differed in their mobility from host lipids. Inferring

that they also were derived from a pool of host lipids, the result indicates the existence of a virus-encoded enzymatic apparatus for modifying host lipids.

The genomic organization of the PSV differs from that of other archaeal viruses. Almost all of the putative genes are found on one strand. Four genes are present on the opposite strand, and the absence of genes on the complementary strand in this region suggests that a genomic inversion may have occurred at the two locations. The genome organization resembles that of some bacterial viruses, for example, T7-like viruses or members of the *Plasmaviridae*, which encode all genes on one strand of the double-stranded DNA genome (van Regenmortel et al., 2000). Nevertheless, the long inverted repeats at the ends of the PSV genome resemble those of the linear genomes of the rudiviruses SIRV1 and SIRV2 of the genus *Sulfolobus*. The sequencing also suggests that, as in case of the rudiviral linear genomes (Blum et al., 2001), the two DNA strands are covalently linked at the genomic termini of PSV.

The most remarkable feature of the PSV genome is that its putative genes yield almost no significant matches with public sequence databases. The best matches are to fragments of three hypothetical proteins encoded in the *P. aerophilum* genome. This contrasts strongly with features of the several double-stranded DNA genomes of hyperthermophilic archaeal viruses, all infecting members of the genera *Sulfolobus* and *Acidianus*, that have previously been sequenced and analyzed. They include circular genomes of the fuselloviruses SSV1 and SSV2 (Palm et al., 1991; Stedman et al., 2003), linear genomes of the lipothrixviruses SIFV (Arnold et al., 2000) and AFV1 (Bettstetter et al., 2003), and rudiviruses SIRV1 and SIRV2 (Peng et al., 2001). In addition, a partial sequence of the lipothrixvirus TTV1 from *T. tenax* has been determined (Neumann, 1988). All these genomes share some homologous genes and yield good sequence matches with genes present in host chromosomes. The latter includes genes encoding glycosyl transferases, CopG family proteins, integrases, dUTPases, and Holliday junction resolvases as well as proteins of unknown function (reviewed in Prangishvili and Garrett, 2004). However, none of the crenarchaeal viruses, including PSV, share clearly homologous genes with viruses of the euryarchaeota or bacteria, and only the rudiviruses share homologous genes with eukaryal poxviruses (Peng et al., 2001).

We propose classifying PSV as the representative of a novel family from its unique morphological features and its exceptional gene content and genome organization. For the new family, we suggest the name *Globuloviridae*, from the Latin *globulus*, a small ball.

Globuloviridae is the fifth novel family to have been introduced over the past few years for viruses of hyperthermophilic archaea, together with the *Fuselloviridae*, *Lipothrixvirudae*, *Rudiviridae*, and *Guttaviridae* (reviewed by Prangishvili et al., 2001). Moreover, the results for the PSV virus reinforce the concept that viruses of hyperthermophilic archaea are exceptionally diverse (reviewed by Prangishvili, 2003). The new evolutionary implications of the properties of the PSV virus, with its unusual morphotype and genome, are still unclear, but elucidating the biology of the virus further will undoubtedly contribute to better understanding of the origin and evolution of viruses in general.

Materials and methods

Enrichment and cloning of cells

The sample from Obsidian Pool, Yellowstone National Park, USA (44°36′35.4″ N and 110°26′20.6″ W) was used to inoculate an 800-ml chemostat culture. Cells were enriched and maintained heterotrophically in the salt base, pH 6, of Allen (1959), supplemented with 0.001% yeast extract, 0.005% peptone, 0.1 mM $CaSO_4{\cdot}2H_2O$, 0.05 mM Na_2SO_4, 0.1 mM KNO_3, and 3 mM $Na_2S_2O_3.5H_2O$. The medium was continuously replaced (dilution rate of 12 ml h^{-1}) and the chemostat was maintained under strictly anaerobic conditions at 85 °C and flushed with N_2 and CO_2 (80/20 v/v, 12 ml h^{-1}). Microbial growth was monitored by directly counting cells using a phase-contrast microscope and a Thoma counting chamber. Cloning of cells from the growing culture was performed using optical tweezers as described by Huber et al. (1995) or by spreading on Gelrite plates.

Cell viability assay

The viability of cells was determined by LIVE/DEAD *Bac*Light Bacterial Viabiliy Kit (Molecular Probes, Leiden, The Netherlands). Experimental protocols followed the instructions of the manufacturer.

Purification of VLPs and the PSV virus

For preparation of VLPs, cells were removed from the enrichment culture by centrifuging for 30 min at 4500 rpm in a Sorvall GS3 rotor and the supernatant was passed through a 0.2-μm filter (Acrodisc PF 0.8/0.2, Pall Gelman Laboratory, Ann Arbor, MI, USA). VLPs were concentrated by ultrafiltration through a 100-kDa PVDF Filter (Pall). Particles were purified from the concentrated suspension by centrifuging in a CsCl buoyant density gradient (0.45 g ml^{-1}) in a Beckman SW60 rotor at 48 000 rpm for 36 h. The fractions were isolated with a syringe, dialyzed against 20 mM Tris-acetate, pH 6, and examined by electron microscopy for the presence of virus particles.

PSV was purified from cell-free cultures of PSV-infected *Pyrobaculum* sp. D11 as described above for the VLPs.

Protein analysis

Proteins were analyzed in SDS-polyacrylamide gels as described by Schägger and Jagow (1987) and stained with Coomassie Brilliant Blue R-250 (Serva). The acrylamide

concentration in the separating gel was 12.5%. Before electrophoresis, virus particles were treated at 95 °C for 5 min with SDS and 2-mercaptoethanol at final concentrations of 2% and 3%, respectively.

Lipid analysis

The cell pellet from 200 ml of an exponentially growing *Pyrobaculum* sp. D11 culture and the purified virus preparation were both dried in a Speed-Vac concentrator (Savant Instruments, Inc. Hicksville, USA), resuspended in 500 µl of chloroform/methanol (1:1), and incubated at 60 °C overnight. Undissolved components were removed by centrifugation at 10 000 × *g*. The supernatants were evaporated and the resulting dried extracts were dissolved in 50 µl chloroform/methanol/H_2O (65/24/4) and briefly incubated at 60 °C. Lipids in the extracts were separated by thin layer chromatography with chloroform/methanol/H_2O (65/24/4) on a Silica gel 60 plate (VWR International-Merck). After the run, the plate was sprayed with 0.5 g $Ce(SO_4)_2 \times 4H_2O$ in 2 N H_2SO_4 and incubated at 160 °C for 15 min.

DNA preparation

Viral particles were disrupted by treatment with 1% SDS for 1 h at room temperature, and extracted with phenol and phenol/chloroform. DNA was precipitated from the aqueous phase by adding sodium acetate (pH 5.3) to a final concentration 0.3 M, and 0.8 vol of isopropanol. The DNA pellet was washed twice with 70% ethanol, air dried, and resuspended in an appropriate volume of buffer TE containing 10 mM Tris–HCl, pH 8, and 1 mM EDTA.

For preparation of *Pyrobaculum* genomic DNA, cells were grown until mid-exponential phase, collected by centrifugation, and suspended in buffer TE. Cells were lysed by addition of SDS to a final concentration of 1% and freezing (−80 °C) and thawing three times. Subsequent steps were performed as described for the viral DNA.

Electron microscopy

Samples were applied to carbon-coated copper grids, negatively stained with 3% uranyl acetate, pH 4.5, and examined using a CM12 transmission electron microscope (FEI, Eindhoven, The Netherlands) operated at 120 keV. Magnification was calibrated using catalase crystals negatively stained with uranyl acetate (Reilein, 1998). Electron micrographs were digitally recorded using a slow-scan CCD camera connected to a PC executing the TVIPS software (TVIPS GmbH, Gauting, Germany).

16S rRNA gene sequence analysis

Isolation of nucleic acid and amplification of 16S rRNA genes were performed as described by Eder et al. (1999). The archaeal sequences were determined using the forward and reverse primers, 8aF, 767aR, 1119aR, and 1512uR (Eder et al., 1999), while bacterial sequences were obtained using primers 519uF and 1406uR (Eder et al., 1999). For the phylogenetic analyses, sequences were aligned with approximately 10 000 homologous sequences available in public databases using the automatic alignment tools of the ARB package (Ludwig and Strunk, 2002). Distance matrix (Jukes–Cantor correction), maximum parsimony, and maximum likelihood (fastDNAml) methods were employed for tree reconstruction using the ARB software.

DNA sequencing and sequence analysis

A shot-gun library was prepared for the PSV viral genome. DNA was sonicated to produce fragments of about 2 kb and these were cloned into the *Sma*I site of the pUC18 vector. DNA was extracted from single colonies using a Model 8000 Robot (Qiagen, Westburg) and sequenced in a MegaBACE 1000 Sequenator (Amersham Biotech). The termini of the viral genome were sequenced by primer walking on viral DNA prepared by the GenomiPhi DNA Amplification Kit (Amersham Biotech). The viral sequence was assembled using the Sequencher Program 3.1. Searches for sequence matches were made in public sequence databases (Altschul et al., 1997; Tatusova and Madden, 1999). Motifs were identified by searching the PFAM database (http://pfam.wustl.edu/cgi-bin/).

Acknowledgments

We thank Brian Hedlund and Werner Reitberger for assistance in enriching environmental samples and purifying strains, Hien Phan for her help in amplifying and sequencing viral DNA, and Harald Huber for help with 16S rDNA sequence analysis. Hans-W. Ackermann is thanked for stimulating discussions and critical reading of the manuscript. Ian Molineux and anonymous reviewers are also thanked for their critical and helpful comments. Samples from Obsidian Pool in Yellowstone National Park were obtained under sampling permit 00115. The work was supported by grants from the Deutsche Forschungsgemeinschaft (PR 663/2-1), the Danish Natural Science Research Council for an Archaea Centre, and the European Union (*Sulfolobus* project grant-QLK3-2000-00649). X.P. received a postdoctoral fellowship from the Danish Technical Science Research Council.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AJ635161, AJ635162, AJ630373.

References

Allen, M.B., 1959. Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. Arch. Microbiol. 32, 270–277.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T., Weidmann, J.F., Kristjanson, J., Klenk, H.P., Nelson, K.E., Fraser, C.M., 2000. A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. Virology 267, 252–266.

Bettstetter, M., Peng, X., Garrett, R.A., Prangishvili, D., 2003. AFV1, a novel virus infecting hyperthermophilic archaea of the genus *Acidianus*. Virology 315, 68–79.

Blum, H., Zillig, W., Mallok, S., Domdey, H., Prangishvili, D., 2001. The linear genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses. Virology 281, 6–9.

Dyall-Smith, M., Tang, S.-L., Bath, C., 2003. Halobacterial viruses: how diverse are they? Res. Microbiol. 154, 309–313.

Eder, W., Ludwig, W., Huber, R., 1999. Novel 16S rRNA gene sequences retrieved from highly saline brine sediments of Kebrit Deep Sea. Arch. Microbiol. 172, 213–218.

Huber, H., Stetter, K.O., 2001. Order I: Thermoproteales. In: Garrity, G. (Ed.), Bergey's Manual for Systematic Bacteriology, 2nd ed., vol. 1. Springer-Verlag, New York, NY, p. 170.

Huber, R., Burggraf, S., Meyer, T., Barns, S.M., Rossnagel, P., Stetter, K.O., 1995. Isolation of a hyperthermophilic archaeum predicted by in situ RNA analysis. Nature 376, 57–58.

Ludwig, W., Strunk, O., 2002. ARB: A software environment for sequence data. Available at http://www.arb-home.de.

Neumann, H., 1988. "Struktur, Funktion und Variabilität eines archaebacteriales Genomes", PhD dissertation, Ludwig-Maximilians-Universität München, Munich, Germany.

Palm, P., Schleper, C., Grampp, B., Yeats, S., McWilliam, P., Reiter, W.-D., Zillig, W., 1991. Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. Virology 185, 2242–2250.

Peng, X., Blum, H., She, Q., Mallok, S., Brügger, K., Garrett, R.A., Zillig, W., Prangishvili, D., 2001. Sequences and replication of genomes of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. Virology 291, 226–234.

Prangishvili, D., Arnold, H.P., Götz, D., Ziese, U., Holz, I., Kristjansson, J.K., Zillig, W., 1999. A novel virus family, the *Rudiviridae*: structure, virus-host interactions and genome variability of the *Sulfolobus* viruses SIRV1 and SIRV2. Genetics 152, 1387–1396.

Prangishvili, D., 2003. Evolutionary insights from studies on viruses from hot habitats. Res. Microbiol. 154, 289–294.

Prangishvili, D., Garrett, R.A., 2004. Exceptionally diverse morphotypes and genomes of crenarchaeal hyperthermophilic viruses. Biochem. Soc. Trans. 32, 204–208.

Prangishvili, D., Zillig, W., 2001. Viruses of the archaea. In: Brenner, S., Miller, J.H. (Eds.), Encyclopedia of Genetics, vol. 4. Academic Press, San Diego, pp. 2114–2116.

Prangishvili, D., Stedman, K.M., Zillig, W., 2001. Viruses of the extremely thermophilic archaeon *Sulfolobus*. Trends Microbiol. 9, 39–42.

Rachel, R., Bettstetter, M., Hedlund, B.P., Häring, M., Kessler, A., Stetter, K.O., Prangishvili, D., 2002. Remarkable morphological diversity of viruses and virus-like particles in terrestrial hot environments. Arch. Virol. 147, 2419–2429.

Reilein, A., 1998. Preparation of catalase crystals. http://www.itg.uiuc.edu/publications/techreports/98-009.

Reiter, W.-D., Palm, P., Zillig, W., 1988. Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. Nucleic Acids Res. 16, 2445–2459.

Rice, G., Stedman, K.M., Snyder, J., Wiedenheft, B., Brumfield, S., Mc Dermott, T., Young, M., 2001. Novel viruses from extreme thermal environments. Proc. Natl. Acad. Sci. U. S. A. 98, 13341–13345.

Schägger, H., Jagow, G., 1987. Tricine dodecyl sulfate-polyacrylamide gel electrophoresis for separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. 166, 368–379.

Slupska, M.M., King, A.G., Fitz-Gibbon, S., Besemer, J., Brodovsky, M., Miller, J.H., 2001. Leaderless transcripts of the crenarchaeal hyperthermophile *Pyrobaculum aerophilum*. J. Mol. Biol. 309, 347–360.

Stedman, K.M., She, Q., Phan, H., Arnold, H.P., Holz, I., Garrett, R.A., Zillig, W., 2003. Biological and genetic relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. Res. Microbiol. 154, 295–302.

Tatusova, T.A., Madden, T.L., 1999. Blast 2 sequences—A new tool for comparing protein and nucleotide sequences. FEMS Microbiol. Lett. 174, 247–250.

Tolstrup, N., Sensen, C.W., Garrett, R.A., Clausen, I.G., 2000. Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*. Extremophiles 4, 175–179.

van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wickner, R.B. (Eds.), 2000. Virus Taxonomy: The Classification and Nomenclature of Viruses. The Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego. 1167 pp.